It is well known that the use of antenna arrays at both sides of communication link can result in high channel capacities provided that the propagation medium is rich scattering. In most previous works presented on MIMO wireless structures, Rayleigh fading conditions were considered. In this work, the capacity of MIMO systems under fully correlated (i.e., correlations between rows and columns of channel matrix) fading is considered. We use replica method and character expansions to calculate the capacity of correlated MIMO channel in closed form. In our calculations, it is assumed that the receiver has perfect channel state information (CSI) but no such information is available at the transmitter.

1. Introduction

The Shannon capacity of a channel defines its theoretical upper bound for maximum rate of data transmission at an arbitrary small bit error probability, without any delay or complexity constraints. Therefore, this capacity represents not only an optimistic upper bound but also a benchmark against which to compare the spectral efficiency of all practical communications systems. In the context of wireless communications systems, initial publication dealt with capacity of (i) single-input single-output (SISO) (no diversity) Rayleigh [7, 13], Rice [12], and Nakagami [21] channels, then (ii) single-input multiple-output (SIMO) (i.e., receiver diversity) Rayleigh [1], and Rice [12] channels. More recently, multiple-input multiple-output (MIMO) has been an important subject of research during the past few years [17, 20]. When there is enough multipath, that is, in a rich scattering environment, the capacity of a MIMO channel can be efficiently multiplied, by adding antenna elements at both transmitter and receiver. MIMO systems are hence a promising solution for high bit rate applications that will provide a breakthrough in the future wireless communication. In a very recent paper [19], Smith and Shafi used central limit argument to show that the capacity of MIMO channels is asymptotically/approximately Gaussian and based on simulations, they found that this approximation was surprisingly accurate for virtually all numbers of transmitter/receiver antenna elements over Rayleigh environment. The implication of this result is that only mean and
variance are needed to fully characterize a useful performance measure of MIMO systems. However, their result is limited to Rayleigh channel.

The highest spectral efficiency of a point-to-point MIMO system is only achieved when there is uncorrelated fading among pairs of transmitter and receiver antennas. In practice, this can be achieved with sufficient spacing among the base station, (mobile or portable) terminal antennas. At the base station, decorrelation is achieved using approximately $10 \lambda$ separation between the nearest elements in a linear array. A mobile station operating in an urban outdoor area or indoor in a home or in office environment is likely to be surrounded by multiple scatterers that contribute to wide-angle scattering. Therefore, only $0.5 \lambda$ spacing maybe adequate. At 2 GHz, this wavelength is 15 cm. This means that even modest four-element antenna array at the base station will have a span of 4.5 m. Thus, contrary to popular belief, the spacing issue is of considerable significance at the base station where mounting of antennas is subject to strict environmental regulations. Note that design and mounting of closely spaced antennas on a small form-factor portable devices (laptops, PDA, etc.) pose additional problems (to just using separation to decorrelate received signals) due to coupling via substrate.

It is of interest to investigate the capacity of multiple-antenna communications systems when there are correlations. In [6, 14], the authors consider capacity when there are correlations at only one end of the system, that is, correlations between rows or columns of the channel matrix, and gives asymptotic results. In [16], the authors give the exact expression for the ergodic capacity of MIMO systems when there is correlation at one end of the system with arbitrary number of antennas. However, their analysis cannot be extended to correlations at both ends (i.e., correlations between rows and between columns of channel matrix) of the communication system. In this paper, we consider the problem for the case in which there are correlations at both ends of the communication system. We assume that the receiver has perfect knowledge of channel state information but the transmitter does not have any knowledge of fading statistics nor instantaneous channel state information. Borrowing the results from [3, 18], we give closed-form expression for the capacity of fully correlated MIMO system.

This paper is organized as follows. Section 2 gives system and channel models. Section 3 gives ergodic capacity of correlated Rayleigh channel. In Section 4, we briefly discuss replica method and evaluate the expression for correlated Rayleigh MIMO capacity. Section 5 briefly explains the direct evaluation of the correlated MIMO capacity. Conclusions are given in the last section.

2. System and channel models

A discrete-time baseband channel model is considered, and flat (nondispersive) fading conditions are assumed and for simplicity, we assume equal number of elements of transmit-receive pairs $(N, N)$. The channel matrix $H$ of dimension $N \times N$ will characterize the channel. The received signal at $j$th antenna is

$$y_j = \sum_{i=1}^{N} h_{ij} x_i + v_j, \quad (2.1)$$
where $y_j$ is the received signal at the receiver antenna $j$, reflecting contributions from all transmit antennas, signals $x_i$, and $v_j$ represents additive white Gaussian noise seen at each receiver antenna. We will assume (without loss of generality) that the noise has unit generalized variance (i.e., $E[vv^T] = I_N$). The signal vector $(N \times 1)$ received at the output can be written as

$$
\begin{pmatrix}
y_1 \\
\vdots \\
y_N
\end{pmatrix} =
\begin{pmatrix}
h_{11} & \cdots & h_{N1} \\
\vdots & \ddots & \vdots \\
h_{1N} & \cdots & h_{NN}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
\vdots \\
x_N
\end{pmatrix} +
\begin{pmatrix}
v_1 \\
\vdots \\
v_N
\end{pmatrix},
$$

or in vector notation, we have

$$y = Hx + v. \quad (2.3)$$

It is assumed that the channel is stationary, ergodic (i.e., time average converges to its ensemble average), and is independent of channel input $x$ and the noise $v$. The channel gains $\{H_{ij}\}$ are complex gains from $j$th transmitter antenna element to the $i$th receiver antenna element.

The input power is constrained to $\rho$, that is, transmit power available at the transmitter,

$$Tr(E[xx^H]) \leq \rho. \quad (2.4)$$

We assume that the receiver has perfect channel state information (CSI) while the transmitter has neither instantaneous nor statistical CSI and as such, the total transmitted power is distributed equally among $N$ transmitting antenna elements (may not be optimal power-allocation scheme when there are correlations), that is,

$$E[xx^H] = \frac{\rho}{N} I_N, \quad (2.5)$$

where $\rho$ is the signal-to-noise ratio (SNR). Our expression can be thought of as lower bound on the capacity of fully correlated MIMO system (owing to the fact that we divide power equally across each antenna, which may not be capacity-achieving scheme). The following notations are used in the paper. $I_N$ represents $N \times N$ identity matrix, $Tr(X)$ is the trace of $X$, $\det Y$ is determinant of $Y$, $\otimes$ denotes the kronecker product.

We will deal exclusively with the above linear model and derive the capacity of MIMO channel models.

### 3. Ergodic capacity of MIMO channels

We assume that $H$ is Gaussian random matrix whose realization is known at the receiver, or equivalently, the channel output consists of the pair $y$, $H$. The input power is distributed equally over all transmitting antennas. Owing to the correlations at both ends, the MIMO channel may be modeled as

$$H = R^{1/2} HS^{1/2}. \quad (3.1)$$
Assuming block-fading model, then the ergodic capacity of a random MIMO channel is given by

\[
C = E_H \left\{ \log \det \left( I_N + \frac{\rho}{N} HH^H \right) \right\},
\]

where \( E_H \) denotes that the expectation is taken with respect to ensemble statistics of \( H \), which is correlated complex Gaussian density in our case.

4. Replica method and evaluation of correlated MIMO capacity

Now we will also briefly introduce a technique found in theoretical physics termed as replica analysis \([9, 22]\). In 1975, Edwards and Anderson, when studying disordered systems of spins, proposed a new method for the investigation of disordered systems—the so-called replica method. In this method, one replaces a single disordered system by \( n \) systems which are identical to the original. Then, for example, instead of calculating the free energy \( F = T \log Z \), one calculates the quantity \( F_n = T \frac{d}{dn} Z^n \). The limit of \( F_n \) as \( n \to 0 \) coincides with the free energy \( F \). Mathematically, this is represented by

\[
F = T \log Z.
\]

The above expression can be calculated as

\[
F = T \left( \lim_{n \to 0} \frac{d}{dn} Z^n \right).
\]

Using the above expression, we can write the capacity expression (3.2) as

\[
C = \lim_{n \to 0} \frac{d}{dn} E_H \left\{ \det \left( I_N + \frac{\rho}{N} HH^H \right)^n \right\},
\]

where the expectation is with respect to

\[
P(H) = C_1 \exp \left( - \text{Tr} \left\{ R^{-1} HS^{-1} H^H \right\} \right),
\]

where the constant \( C_1 \) is given by

\[
C_1 = \frac{\pi^{-N^2}}{\det(R)^N \det(S)^N}.
\]

Using replica method, we can write

\[
C = \lim_{n \to 0} \frac{d}{dn} E_H \left\{ \det \left( I + \alpha HH^H \right)^n \right\},
\]

where we have defined \( \alpha = \rho/N \). Now we define

\[
G = E_H \left\{ \det \left( I + \alpha HH^H \right)^n \right\}.
\]
Performing the eigendecomposition of matrix $HH^H$, we get

$$G = C_1 C_2 \int \prod_{i=1}^{N} d\lambda_i (1 + \alpha \lambda_i)^\mu \Delta(\lambda)^2 Z(\lambda), \quad (4.8)$$

where $C_2$ is normalization constant of the transformation and is given in [16]. In the above equation, we have that $\lambda_i$ are the eigenvalues of $HH^H$, $\Delta(\lambda)^2 = \prod_{i<j} (\lambda_i - \lambda_j)^2$ is the square of the Vandermonde determinant, which is Jacobian of transformation from $HH^H \rightarrow U\Sigma U^H$. $Z(\lambda)$ is given by

$$Z(\lambda) = \int dU \int dV \exp \left( - \text{Tr} \{ R^{-1} UhVS^{-1} V^H h^H U^H \} \right), \quad (4.9)$$

where $h$ are the singular values of matrix $H$ and $U$, $V$ are unitary matrices which are integrated with the Haar measure over $U(N)$ (unitary group). It is clear from the expression for $Z(\lambda)$ that this integral cannot be solved using well-known HIZ (Harish-Chandra-Itzykson-Zuber) formula, which would have been possible if either matrix $R$ or $S$ is multiple of identity matrix. To tackle the integral $Z(\lambda)$, we employ character expansion method. Character expansion has been used successfully to reduce the number of degrees of freedom from $N^2$ to $N$ in matrix models and solve the resulting problem by saddle-point method (for large systems) [10, 11]. Group characters have been used by number of researchers to solve mathematical problems. In this paper, we follow [3, 18]. In character expansion method, $\exp (- \text{Tr} \{ R^{-1} UhVS^{-1} V^H h^H U^H \})$ is expanded as an invariant function of variable $R^{-1} UhVS^{-1} V^H h^H U^H$ of irreducible representation of $r$ of the general linear group (GL($N$) group),

$$\exp (- \text{Tr} \{ R^{-1} UhVS^{-1} V^H h^H U^H \}) = \sum_r f_r \chi_r (R^{-1} UhVS^{-1} V^H h^H U^H), \quad (4.10)$$

where the coefficients are functions of the $N$ highest weight components of the representation

$$r = \{0 \leq k_N \leq \cdots \leq k_1 < \infty \}. \quad (4.11)$$

$\sum_r$ is nothing but is the sum over $N$ ordered integers ($k_j = N + n_j - j$) [3]. $n_j$ is the partition of a $U(N)$ group into $N$ parts ($n_1, \ldots, n_N$), where $n_1 \geq \cdots \geq n_N$. A character is just the trace taken in the proper representation, that is, $r$ in our case, so we have

$$\chi_r (R^{-1} UhVS^{-1} V^H h^H U^H) = R_r^{ab} U_r^{bc} h_r^{cd} V_r^{de} S_r^{ef} V_r^{fg} h_r^{gh} U_r^{*}. \quad (4.12)$$

We now state the orthogonality property [3, 5, 8]

$$\int dU U_{rs}^r U_{cd}^s = \frac{1}{d_r} \delta_{rs} \delta_{ac} \delta_{bd}, \quad (4.13)$$

where $U^r$ and $U^s$ are the group matrix elements in the representations $r$ and $s$ respectively, $d_r$ is the dimension of the representation, and $\delta(\cdot)$ is Kronecker delta. Combining
(4.9), (4.12), and (4.13), we can write

\[ Z(\lambda) = \sum_r \frac{f_r}{d_r^2} \chi_r(R^{-1}) \chi_r(S^{-1}) \chi_r(\lambda). \]  

(4.14)

Now we evaluate the term \( \frac{f_r}{d_r^2} \) in the above equation. To do so, we observe that the author in [3] found that

\[ f_r = s(k) \det \left( \frac{1}{(k_j + i - N)!} \right) = \frac{s(k) \Delta(k)}{C(k)}, \]  

(4.15)

where \( C(k) = \prod_{j=1}^{N} k_j! \). \( s(k) = (-1)^v \) with \( v = N(N-1)/2 - \sum_{j=1}^{N} k_j \). In [3], it is also shown that

\[ \frac{f_r}{d_r} = \frac{s(k) F_N}{C(k)} \]  

(4.16)

with \( F_N = \prod_{j=1}^{N-1} j! \). From (4.15) and (4.16), it is straightforward to obtain

\[ \frac{f_r}{d_r^2} = \frac{s(k) F_N^2}{\Delta(k) C(k)}. \]  

(4.17)

A character is defined by Weyl’s classical formula [4, 10, 11]

\[ \chi_r(A) = \frac{\det \left( \lambda_{i,j}^{k} \right)_{i,j=1,...,N}}{\Delta(a)}, \]  

(4.18)

where \( a_i \) are the eigenvalues of matrix \( A \). \( \Delta(a) \) is the Vandermonde determinant. The expression \( \det(a_{i,j}^{k})_{i,j=1,...,N} \) is the generalized Vandermonde determinant. An alternative form for the character formula is given by [3]

\[ \chi(A) = \det(y_{n,i+j}), \]  

(4.19)

where \( y_n \) is the complete symmetric function in the argument \( t_1, \ldots, t_N \) of degree \( n \).

Substituting (4.18) into (4.14), we get

\[ Z(\lambda) = \frac{1}{\Delta(\lambda)} \sum_r \beta_r \det \left( \lambda_{i,j}^{k} \right)_{i,j=1,...,N}, \]  

(4.20)

where \( \beta_r \) is given by

\[ \beta_r = \frac{f_r}{d_r^2} \chi_r(R^{-1}) \chi_r(S^{-1}). \]  

(4.21)
The Vandermonde determinant can be written as [5]

$$\Delta(\lambda) = \prod_{i<j}(x_i - x_j) = \det \begin{pmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_N \\ \vdots & \vdots & \vdots \\ \lambda_1^{N-1} & \cdots & \lambda_N^{N-1} \end{pmatrix}. \tag{4.22}$$

Vandermonde determinant can also be expressed in terms of orthogonal polynomial. We make use of the orthogonal polynomial $P_n(\lambda)$, (orthogonal with respect to the measure or weight) [5]

$$\int_0^\infty d\lambda e^{-V(\lambda)}P_n(\lambda)P_m(\lambda) = h_n \delta_{nm}, \tag{4.23}$$

where $\delta_{nm}$ is Kronecker delta, and $h_n$ is normalization constant. The $P_n$s are orthogonal polynomials and are functions of single real variable $\lambda$. If $P_m(\lambda) = \lambda^m + \sum_{j=0}^{m-1} P_{m,j}\lambda^j$ are monic polynomials of degree $m$, for $m = 0, 1, \ldots, N - 1$, then

$$\Delta(\lambda) = \det(\lambda_i^{j-1})_{1 \leq i, j \leq N} = \det(P_{i-1}(\lambda_j))_{1 \leq i, j \leq N}. \tag{4.24}$$

This is easily proved by performing suitable linear combinations of columns, a procedure that leaves the determinant unchanged. The determinant is indexed with $i, j \in \{1, 2, \ldots, N\}$. Hence, the Vandermonde determinant can be written as (in terms of orthogonal polynomial)

$$\Delta(\lambda) = \det \begin{pmatrix} P_0(\lambda_1) & \cdots & P_0(\lambda_N) \\ P_1(\lambda_1) & \cdots & P_1(\lambda_N) \\ \vdots & \vdots & \vdots \\ P_{N-1}(\lambda_1) & \cdots & P_{N-1}(\lambda_N) \end{pmatrix}. \tag{4.25}$$

Using (4.20), (4.21), and (4.24), we can further write (4.8) as

$$G = C_1 C_2 \int_0^\infty \prod_{i=1}^N d\lambda_i (1 + \alpha \lambda_i)^n \sum_r \beta_r \det(\phi_i^{k_j}) \det(\lambda_i^{j-1}). \tag{4.26}$$

Let $C_3 = C_1 C_2$, the above equation can further be written as

$$G = C_3 \int_0^\infty \prod_{i=1}^N d\lambda_i (1 + \alpha \lambda_i)^n \sum_r \beta_r \det(\phi_i(\lambda_i)) \det(\lambda_i^{j-1}), \tag{4.27}$$

where we have set $\det(\phi_i^{k_j}) = \det(\phi_j(\lambda_i))$. Here, we need the following formula. Let $\phi_1(k), \ldots, \phi_N(k)$ and $\psi_1(k), \ldots, \psi_N(k)$ be arbitrary functions in $k$. Then [15],

$$\det(\phi_i(k_i)) \det(\psi_j(k_j)) = \sum_{\sigma \in S_n} \det(\phi_i(k_{\sigma(j)})) \psi_j(k_{\sigma(j)}), \tag{4.28}$$
where $\sigma$ runs over all permutations of $S_N$. The above formula can be proved straightforwardly by calculating the left-hand side by the usual product formula of the determinant. Then it becomes a summation of $N^N$ terms. Because of the multilinearity of the determinants, only $N!$ of these terms are nonzero. Rearranging the $N!$ terms, we obtain the required formula. Using the above expression, we can write

\[
G = C_3 \int_0^\infty \prod_{i=1}^N d\lambda_i (1 + \alpha\lambda_i)^n \sum_r \beta_r \sum_{\sigma \in S_N} \det(\phi_j(\lambda_{\sigma(i)})\lambda_{\sigma(i)}^{i-1})
\]

Plugging the above equation into (4.3) and taking the replica limit give the ergodic capacity of MIMO system in fully correlated system (correlations at both ends). After writing the paper, the authors came to know about Cauchy-Binet theorem [2] which can be applied directly to (4.27) to get the final answer, that is, (4.30). The Cauchy-Binet theorem is given:

\[
\int \prod_{i=1}^N d\mu(\lambda_i) \det(A_k(\lambda_i)) \det(B_k(\lambda_i)) = N! \det \left( \int d\mu(\lambda) A_k(\lambda) B_k(\lambda) \right),
\] (4.31)

where $d\mu(\lambda_i)$ is the integration measure.

5. Alternative expression for correlated MIMO capacity

The channel capacity can also be calculated without the help of replica analysis, that is, it can be calculated by direct calculation. The channel capacity for correlated MIMO system is given by (3.2) with probability distribution given in (4.4). Performing the eigendecomposition of $HH^H$, we get

\[
C = C_1 C_2 \int_0^\infty \prod_{i=1}^N d\lambda_i \sum_{i=1}^N \log (1 + \alpha\lambda_i) \Delta(\lambda)^2 Z(\lambda),
\] (5.1)

where $\lambda_i$, $\alpha$, $C_1$, $C_2$, and $Z(\lambda)$ are the same as in the previous section.

In Figure 5.1, we have capacity results for $N = M = 2$ systems at SNR = 20 dB (signal-to-noise ratio) versus the correlation coefficient. It is clear from the figure that the correlation degrades the capacity of the system. In Figure 5.2, we have shown capacity results
Figure 5.1. Capacity results for \( N = 2 \) systems at SNR = 20 dB. The solid line represents the capacity when there are correlations at the transmitting side as well as the receiving side with the same correlation coefficient matrix. Dashed line is the capacity when there is correlation only at transmitting side.

Figure 5.2. Capacity results for \( N = 2 \). The solid line is capacity when there is no correlation between antennas. Dashed line represents capacity when there is correlation of 0.9 at transmitting side and 0.1 at the receiving side.

for \( N = M = 2 \). It is clear from the figure that as the SNR goes high, the difference between the capacity without correlation and the capacity with high correlation becomes large.

6. Conclusion

In this paper, we derived an expression for the capacity of fully correlated MIMO channel. The expression was obtained by using replica trick and a powerful technique in group
Capacity of MIMO system

theory, known as character expansions. In the case of semicorrelated case (correlations on one side of the communications systems), we can solve the problem without using character expansion, as is given in [16].

References


Ejaz Khan: Department of Electronic and Electrical Engineering, Faculty of Engineering and Architecture, University College Dublin, Belfield, Dublin 4, Ireland
*E-mail address:* ejaz.khan@ee.ucd.ie

Conor Heneghan: Department of Electronic and Electrical Engineering, Faculty of Engineering and Architecture, University College Dublin, Belfield, Dublin 4, Ireland
*E-mail address:* conor.heneghan@ucd.ie
Submit your manuscripts at http://www.hindawi.com